Human Molecular Genetics, 2000, Vol. 9, No. 4 489–501

ARTICLE

Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis

Tamim H. Shaikh[1,+], Hiroki Kurahashi[1,+], Sulagna C. Saitta[1], Anna Mizrahy O'Hare[1], Ping Hu[4], Bruce A. Roe[4], Deborah A. Driscoll[1,2,3], Donna M. McDonald-McGinn[1], Elaine H. Zackai[1,2], Marcia L. Budarf[1,2] and Beverly S. Emanuel[1,2,§]

[1]Division of Human Genetics and Molecular Biology, The Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA, Departments of [2]Pediatrics and [3]Obstetrics and Gynecology, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA and [4]Department of Chemistry and Biochemistry, University of Oklahoma, Norman, OK 73019, USA

Received 25 October 1999; Revised and Accepted 4 January 2000

The 22q11.2 deletion syndrome, which includes DiGeorge and velocardiofacial syndromes (DGS/VCFS), is the most common microdeletion syndrome. The majority of deleted patients share a common 3 Mb hemizygous deletion of 22q11.2. The remaining patients include those who have smaller deletions that are nested within the 3 Mb typically deleted region (TDR) and a few with rare deletions that have no overlap with the TDR. The identification of chromosome 22-specific duplicated sequences or low copy repeats (LCRs) near the end-points of the 3 Mb TDR has led to the hypothesis that they mediate deletions of 22q11.2. The entire 3 Mb TDR has been sequenced, permitting detailed investigation of the LCRs and their involvement in the 22q11.2 deletions. Sequence analysis has identified four LCRs within the 3 Mb TDR. Although the LCRs differ in content and organization of shared modules, those modules that are common between them share 97–98% sequence identity with one another. By fluorescence *in situ* hybridization (FISH) analysis, the end-points of four variant 22q11.2 deletions appear to localize to the LCRs. Pulsed-field gel electrophoresis and Southern hybridization have been used to identify rearranged junction fragments from three variant deletions. Analysis of junction fragments by PCR and sequencing of the PCR products implicate the LCRs directly in the formation of 22q11.2 deletions. The evolutionary origin of the duplications on chromosome 22 has been assessed by FISH analysis of non-human primates. Multiple signals in Old World monkeys suggest that the duplication events may have occurred at least 20–25 million years ago.

INTRODUCTION

Non-random acquired and somatic chromosomal changes have been demonstrated in association with numerous human diseases. Recurrent constitutional rearrangements such as translocations, inversions and deletions suggest that there may be preferred chromosomal sites for recombination or rearrangement in the human genome (1–7). Although chromosome 22 represents only 2% of the haploid human genome (8), recurrent acquired and somatic rearrangements of this chromosome are associated with multiple malignant diseases and developmental abnormalities (reviewed in ref. 9). The majority of these recurrent rearrangements take place within 22q11.2, suggesting genomic instability related to the structure of this region of human chromosome 22.

The non-random chromosome 22 abnormalities include acquired tumor-associated rearrangements such as the t(9;22) associated with acute lymphocytic leukemia and chronic myeloid leukemia, the t(8;22) variant translocation associated with Burkitt's lymphoma and the t(11;22) of Ewing's sarcoma and peripheral neuroepithelioma. The recurrent constitutional abnormalities of 22q include the duplications associated with the supernumerary bisatellited marker chromosome of Cat Eye syndrome (CES) (10), the translocations which give rise to the recurrent t(11;22) malsegregation-derived Supernumerary der(22)t(11;22) syndrome (2,3,11), and the translocations and deletions associated with DiGeorge, velocardiofacial and conotruncal anomaly face syndromes (DGS/VCFS/CAFS) (12–19). Recent studies have demonstrated that deletions of chromosome 22q11.2 occur at a high frequency, estimated at ~1

[+]These authors contributed equally to this work
[§]To whom correspondence should be addressed at: The Children's Hospital of Philadelphia, 1002 Abramson Research Center, 3516 Civic Center Boulevard, Philadelphia, PA 19104, USA. Tel: +1 215 5903856; Fax: +1 215 590 3764; Email: beverly@mail.med.upenn.edu

in 4000 live births (20). The overwhelming majority of deletions occur sporadically as *de novo* lesions, indicating a high 'mutation' rate within this genomic region (~2.5×10^{-4}).

Analysis of patients with DGS, VCFS or CAFS has demonstrated that 85–90% are deleted for the same chromosomal region of 22q11 (17,19,21,22). Of the patients with a deletion the vast majority share a large (>3 Mb) deletion that is flanked by markers *D22S427* in 22q11.21 and *D22S801* within q11.23 (23,24). A smaller 1.5–2 Mb variant deletion occurs in 10–12% of the 22q11 deletion patients (21,23,25). Further, several reports have described patients with unique deletions that are either nested within the large 3 Mb typically deleted region (TDR) (26–30) or have no overlap with the TDR (31,32). However, these unique deletions are rare.

The identification of chromosome-specific low copy DNA repeat (LCR) elements on chromosome 22 has led to the suggestion that these repeats may be responsible for the instability of 22q11 (33–35). Recent evidence indicates that these repeat elements are components of larger (>100 kb), chromosome 22-specific blocks of duplicated sequence (35–37). Copies of the chromosome 22-specific LCRs have been reported at or near the end-points of the typical 3 Mb DGS/VCFS/CAFS deletion on 22q11.2 (23,36) and at the end-points of the CES duplication (24,38). This has led to the hypothesis that homologous recombination between copies of the chromosome 22-specific LCRs mediate the deletions and duplications associated with DGS/VCFS/CAFS and CES as well as other rearrangements of chromosome 22 (23,24,38).

In an effort to better characterize the chromosome 22-specific LCRs, we have constructed a cosmid, bacterial artificial chromosome (BAC) and P1-derived artificial chromosome (PAC) contig between markers *D22S427* and *D22S801*. Using these reagents for fluorescence *in situ* hybridization (FISH) analysis of patient samples, we have identified a 1.5–2 Mb recurrent variant deletion which is distinct from the previously described smaller variant deletion (21,24,25,39). Analysis of the sequence of the entire 3 Mb TDR reveals four LCRs. In this report, we delineate the complicated structure of the four LCRs and demonstrate their involvement in both of the recurrent and in one unique smaller 22q11.2 deletions. These data now provide the first direct evidence that portions of the LCRs on chromosome 22 are contained within deletion-mediated junction fragments of 22q11 and that portions are lost as a result of the rearrangement. Further, although these duplicated sequences have not been identified in the rodent genome, their evolutionary origin has been further assessed. Here we demonstrate duplication events in non-human primate genomes (chimpanzee, gorilla and rhesus) in regions orthologous to human 22q11, suggesting a primate-specific origin and amplification of the LCRs.

RESULTS

Identification of novel but recurrent deletion end-points within the 3 Mb TDR

Routine diagnosis to determine the presence of a 22q11.2 deletion in patients with features of DGS or VCFS was performed utilizing the N25 probe (ONCOR/VYSIS). Two hundred patients with deletions in the TDR were selected for additional study, of which 199 were deleted for N25. The one remaining patient carried a clinical diagnosis of VCFS but was not deleted for N25. To determine the relative position of deletion end-points these 200 patients were further analyzed by FISH using pairwise probe combinations. The probes used for FISH in this analysis were cosmids which contained the markers *D22S427*, *D22S36* (pH11), *HCF2* and *D22S801* (LN80), respectively (Figs 1 and 2). Dual-color pairwise analyses were performed with cosmids for two marker pairs: (i) c103a2 (*D22S36*) with c45c9 (*D22S801*); and (ii) c106e4 (*D22S427*) with c2c9 (*HCF2*) such that each member of a test pair was labeled with a different color. In each experiment, cos82 (*D22S39*), a control cosmid probe was used to mark the telomeric end of chromosome 22 (22q13.3) and as a hybridization control. All cosmids used as FISH probes to define the deletion boundaries were single copy and did not contain duplicated sequences.

The vast majority of the patients (175 of 200, or 87%) have the common large 3 Mb deletion. The deleted chromosome 22 in these 175 patients is positive for both c106e4 and c45c9 (Fig. 1a and b) and negative for both c103a2 and c2c9 (Fig. 1a and b). Therefore, their deletion interval (Fig. 2, A–D region) is flanked proximally by *D22S427* and distally by *D22S801*. Twenty (10%) of the N25-deleted patients were found to have signal for c2c9 (*HCF2*) on both of their chromosomes 22 (Fig. 1c) indicating that these individuals have smaller deletions with a more centromeric distal deletion end-point.

To further define the distal end-point of their deletion, these 20 patients were further analyzed by FISH with c87f9 (*ZNF74*). In 16 of these individuals neither chromosome 22 was deleted for c87f9 (Fig. 1d). FISH of metaphase chromosomes from the 16 patients not deleted for c87f9 was performed with a more centromeric marker, c68a1 (*D22S788*). This demonstrated that all are deleted at this locus (Fig. 1f) suggesting an A–B deletion interval (Fig. 2). Therefore, these 16 individuals have a distal deletion breakpoint which maps between *D22S788* and *ZNF74* similar to patients described previously (21,25). Interestingly, the remaining four individuals were found to be deleted for c87f9 on one of their chromosomes 22 (Fig. 1e). These four individuals appear to have the deletion delineated as A–C (Fig. 2). These data further suggest that the 1.5–2 Mb variant deletion actually consists of two distinct but recurrent deletions whose distal breakpoints differ from one another. Thus, we have identified a novel recurrent distal deletion end-point which is located between markers *ZNF74* and *HCF2* within the 3 Mb TDR.

With regard to the five remaining deleted patients, four appear to have the same variant deletion interval. When examined by FISH with the aforementioned probes these four individuals were not deleted for c103a2 (*D22S36*). These individuals were tested with c87h3 which maps ~50 kb distal to c103a2 (Fig. 2). All four individuals are deleted for c87h3 suggesting that their proximal deletion end-point maps distal to *D22S36* in the region denoted by the small hatched box (Fig. 2). Their distal deletion end-point is flanked by *D22S801*, similar to the distal end-point of the 3 Mb TDR. Thus, this rearrangement appears to be another novel but recurrent deletion whose proximal end-point is located between markers *D22S36* and *D22S75*. The remaining patient is the one with the unique deletion that does not include the N25 probe. This individual (CH98-205) was deleted for *D22S941* (c102g9 in Fig. 2), a marker that is ~100 kb distal to N25. Another probe, c19d3 (Fig. 2), which is ~50 kb proximal to c102g9 was not deleted suggesting that her proximal deletion end-point maps to an ~50 kb region between probes c19d3 and c102g9. Similar to the vast majority of deleted patients, her distal deletion end-point is

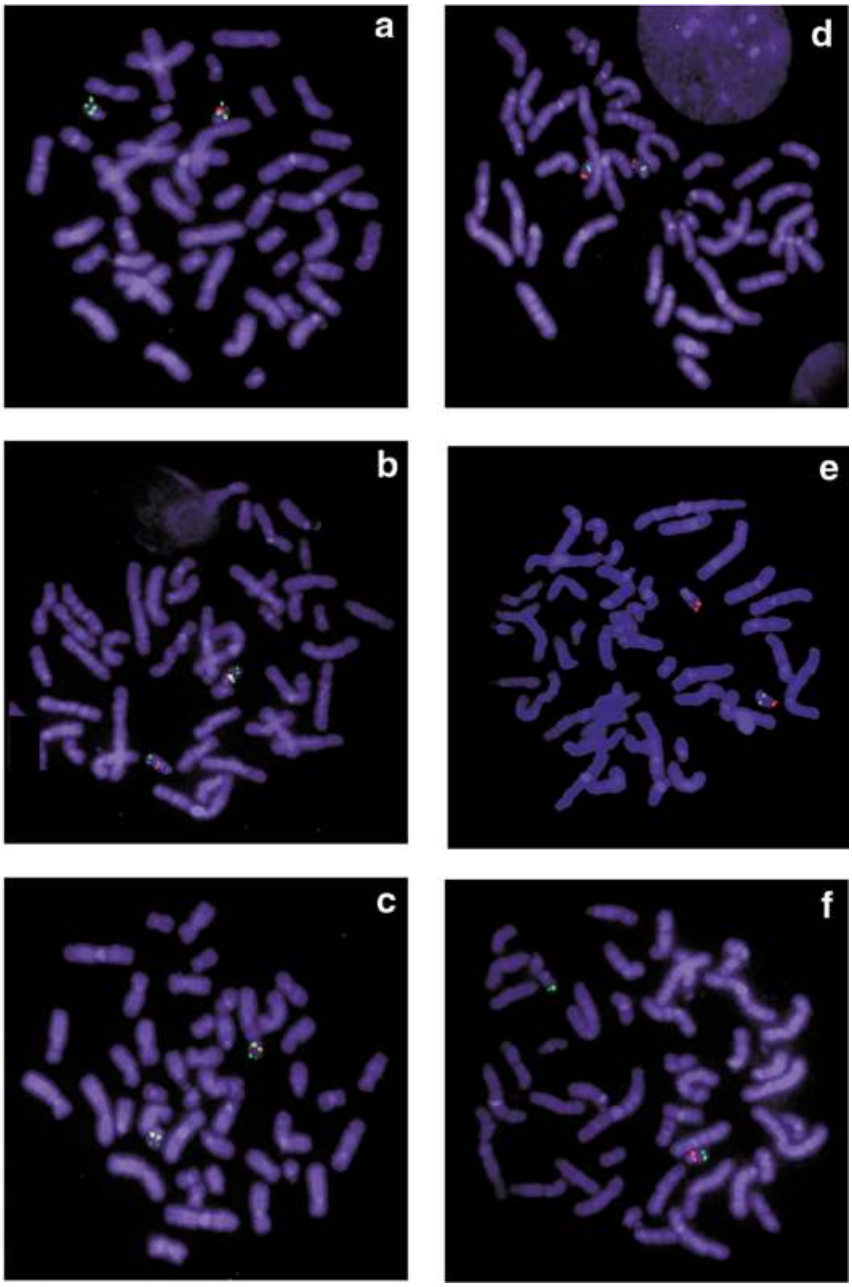

Figure 1. Sizing of 22q11 deletions by dual-color FISH. Metaphase spreads from 22q11.2 deletion patients hybridized with cosmid probes labeled with digoxygenin and detected with rhodamine (red signal) or labeled with biotin and detected with FITC–avidin (green). Standard 3 Mb (A–D) deletion cohybridized with three cosmids. (**a**) c103a2 (red), c45c9 (green) and control probe cos82 (green). (**b**) c106e4 (red), c2c9 (green) and control probe cos82 (green). (**c**) Variant (A–B) deletion cohybridized with three cosmids: c106e4 (red), c2c9 (green) and control probe cos82 (green). (**d**) Variant (A–B) deletion cohybridized with two cosmids: c87f9 (green) and control probe cos82 (red). (**e**) Variant (A–C) deletion cohybridized with two cosmids: c87f9 (green) and control probe cos82 (red). (**f**) Variant (A–B) deletion cohybridized with two cosmids: c68a1 (red) and control probe cos82 (green).

flanked by *D22S801*. In summary, we have identified three recurrent smaller deletions nested within the typically deleted region as well as one unique deletion.

Construction and sequencing of a contig across the 3 Mb TDR

To further characterize the mechanism of deletion and the etiology of the phenotype, a BAC/PAC/cosmid contig across the 3 Mb TDR was constructed between markers *D22S427* and *D22S801* (Fig. 2). The entire contig has been sequenced at the University of Oklahoma (http://www.genome.ou.edu/) (40). This contig includes our previously constructed 1.2 Mb cosmid contig across the DiGeorge chromosomal region (DGCR) between markers *D22S36* and *D22S788* (M.L. Budarf, I.S. Emanuel and B.A. Roe, unpublished data). The sequence of this 1.2 Mb region has provided a useful resource toward the identification of 20–25 genes within this interval (41). To expand the 1.2 Mb DGCR contig, chromosome walks were initiated using PCR-derived

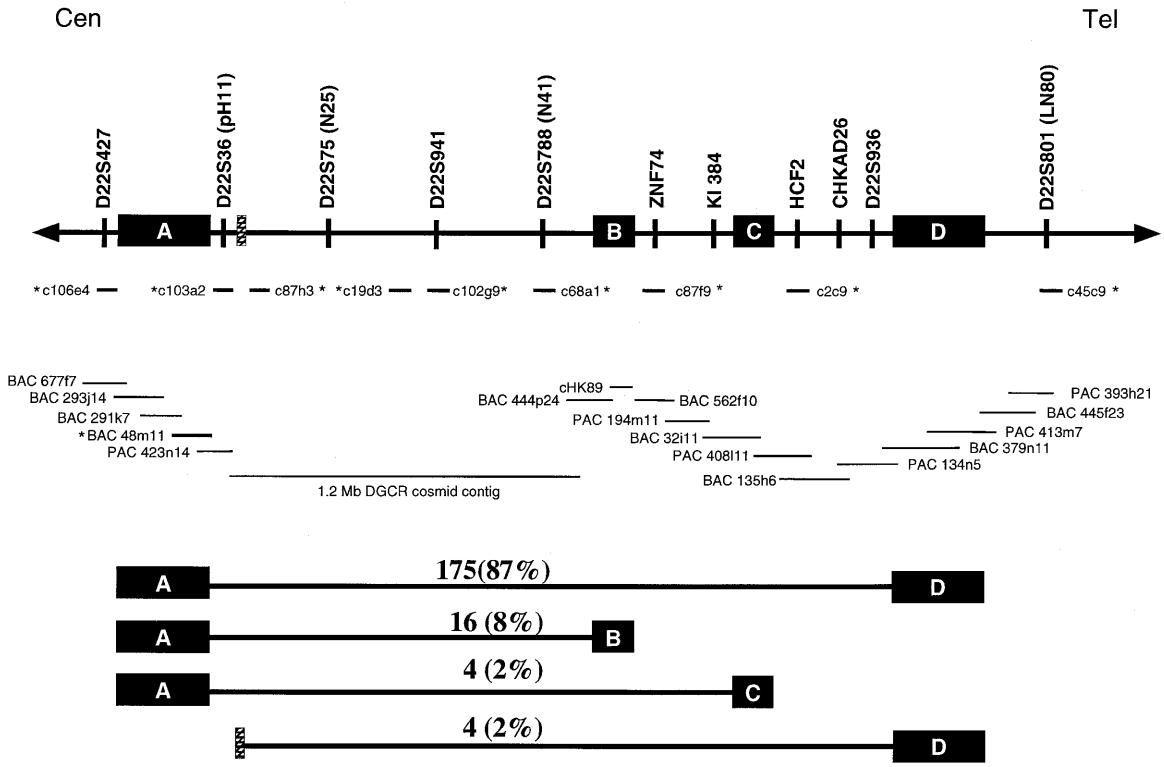

Figure 2. Contig across the 3 Mb TDR. A BAC/PAC/cosmid contig across the entire 3 Mb region between markers *D22S427* and *D22S801* (LN80) is shown. The orientation of the contig is centromere (Cen) to telomere (Tel). LCR-A, -B, -C and -D are indicated as filled boxes. The small hatched box between markers *D22S36* and *D22S75* indicates the proximal deletion end-point of the variant deletion not deleted for c103a2. Markers that anchor the contig are indicated above the line at the top of the figure. The 1.2 Mb DGCR cosmid contig is shown as one single line anchored between markers *D22S36* (pH11) and *D22S788* (N41). The cosmids and BAC used as FISH probes are indicated by asterisks. The numbers above the lines which indicate deletion size (at the bottom of the figure) represent the number of patients identified with those deletion boundaries.

probes from the clones at each end to screen BAC and PAC libraries. Probes derived from unique sequence-tagged site (STS) markers in the region (42) also were utilized. Clones that were positive for the STS markers then were end-sequenced and this sequence was utilized in designing new PCR-derived probes from each end of the clone. These new end probes were used to rescreen the filters to continue the process of chromosome walking in both directions until all the gaps were closed. We have recently isolated a clone cHK89, which extends into the gap between BAC 444p24 and BAC 562f10 from the end of BAC 444p24. cHK89 appears also to overlap with the end of BAC 562f10 by sequence analysis. This gap has been difficult to fill as the region between clones BAC 444p24 and BAC 562f10 is under-represented in the bacterial and yeast libraries screened previously (39,43).

Identification of four chromosome 22-specific LCRs within the 3 Mb TDR

Initial screening with some of our probes yielded a significantly higher than expected representation of clones in the cosmid, BAC and PAC libraries. Further analysis of these clones by Southern hybridization and sequencing confirmed that they contain previously described chromosome 22-specific LCRs (24,33–36,40,44). Due to the presence of these sequence duplications within the clones, confirmation of this contig required extensive sequence analysis. Clones were anchored within unique sequence whenever possible. Clones resulting from chromosome walks with end probes, some of which were duplicated, were confirmed to be contiguous by sequence analysis of multiple markers present within putative overlapping clones. Only clones that shared 100% sequence homology at multiple markers were considered to be truly overlapping. By analysis of the sequence of this contig it was determined that there are four regions within the 3 Mb TDR that contain the chromosome 22-specific LCRs. Here we have designated these four copies of the LCRs A, B, C and D (Fig. 2) and they correspond to LCR22-2, -3a, -3b and -4, respectively (40). Thus, the mapping data (Fig. 2) taken together with the FISH data derived from the analysis of patient samples suggest that the different deletion end-points localize in the vicinity of the chromosome 22-specific LCRs (Fig. 2)

Sequence analysis of LCRs reveals a complex organization

Sequence analysis of clones from LCR-A, -B, -C and -D suggests a complex organization of duplicated modules (Fig. 3). There are differences in the size, content and organization of duplicated modules within each of the LCRs. LCR-A and -D, which correspond to the end-points of the most frequent 3 Mb deletion, appear to be the largest in size and most similar to one another. Both LCR-A and -D contain previously described markers which include *BCRL*, *HMPLPL* (*POM121L*), *GGTL*, *GGTrelL*, *V7rel* and others (23,36). The marker *GGTrelL*, duplicated within LCR-A and -D only, contains the first ~500 bp of the *GGTrel* gene (45).

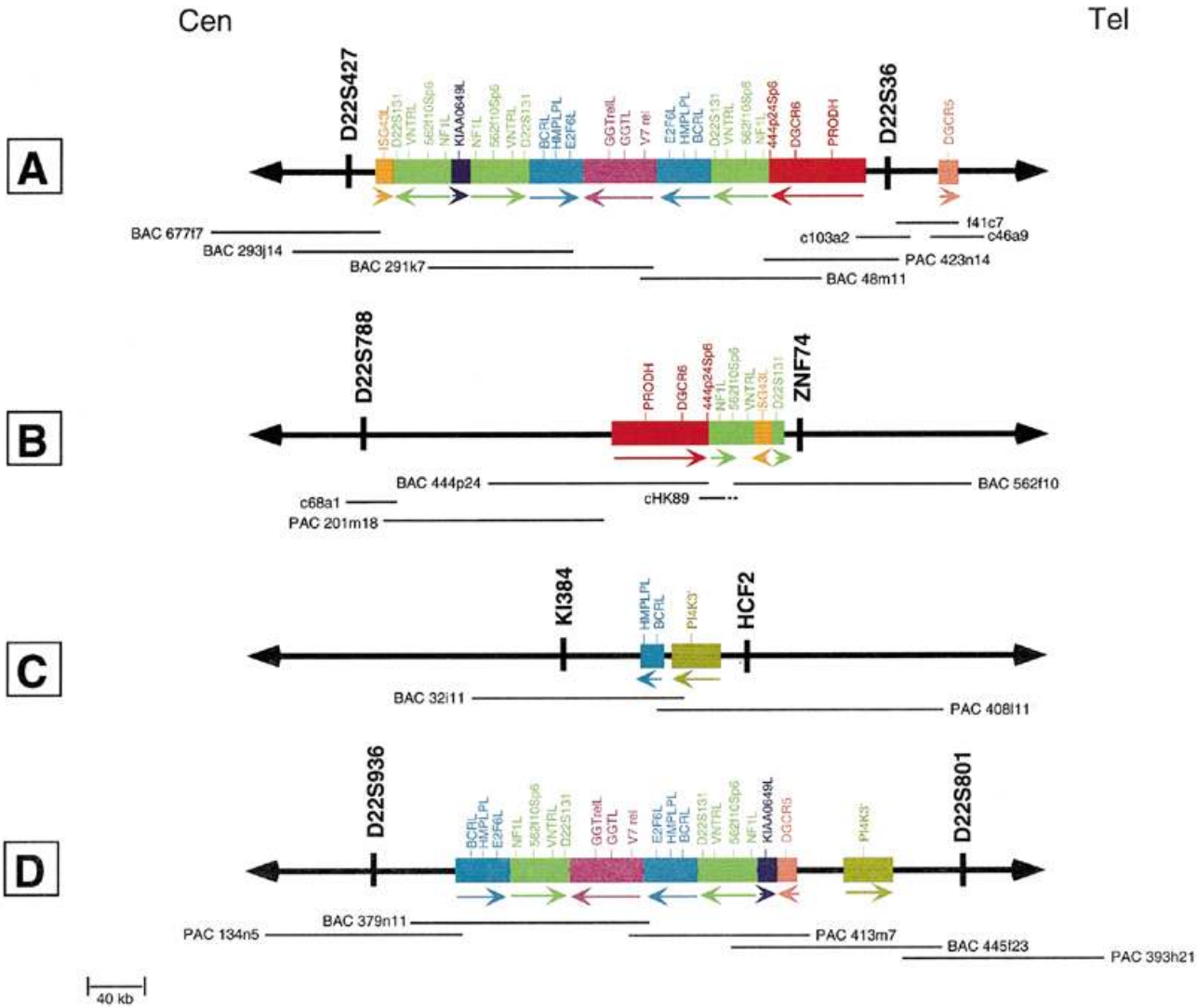

Figure 3. Organization of LCRs. The spatial arrangement of duplicated modules within LCR-A, -B, -C and -D is shown. The duplicated modules are shown as colored boxes and the markers within them are shown above in the same color as the boxes. The orientation of each LCR is centromere to telomere. The sizes of the boxes are proportional to the estimated size of the respective module. Arrows below the duplicated modules indicate their orientation with respect to other copies within the same LCR as well as other LCRs. The BAC/PAC/cosmid contig spanning each LCR is shown below each block. Unique markers flanking the LCRs are shown in black.

In addition to the above markers, the AT-rich sequence shown to be involved in a balanced t(17;22) translocation causing neurofibromatosis type 1 (*NF1-like*) (46) is duplicated within LCR-A, -B and -D. Other markers identified within the LCRs include a VNTR-like repeat which is duplicated within LCR-A, -B and -D (H. Kurahashi, T.H. Shaikh and B.S. Emanuel, unpublished data) and an *E2F6-like* (*E2F6L*) sequence which is duplicated in LCR-A and -D. *E2F6L* appears to be a processed pseudogene of *E2F-6* which is a member of the E2F family of transcription factors (47). Also duplicated within LCR-A and -D is a marker related to *KIAA0649* (*KIAA0649L*) which is a predicted gene previously isolated from a human brain cDNA library (48).

The extent of overlap between the clones which span LCR-A and -D has been determined and their sequence content has been analyzed in detail. Based on this analysis, LCR-A is estimated to be ~350 kb and LCR-D is estimated to be ~250 kb. Sequence data suggest that LCR-A contains two copies of the ~40 kb module containing markers *BCRL*, *HMPLPL* and *E2F6L* (Fig. 3A, blue boxes). LCR-A also contains one copy of an ~75 kb module containing markers 444p24Sp6, *DGCR6* (49) and *PRODH* (proline dehydrogenase) (50) which is duplicated in LCR-B (Fig. 3A and B, red boxes). This module previously has been referred to as sc11.1 (33,34,39). In addition to these modules, LCR-A also contains three copies of an ~45 kb module containing markers *D22S131*, VNTRL, 562f10Sp6 and *NF1L* (Fig. 3A, green boxes), one copy of the ~55 kb module containing markers *GGTrelL*, *GGTL* and *V7rel* and one copy of an ~15 kb module containing marker *KIAA0649L*.

Within LCR-D there are only two copies of the ~45 kb module containing markers *D22S131*, VNTRL, 562f10Sp6 and *NF1L*. Further, the ~75 kb module containing markers 444p24Sp6, *DGCR6* and *PRODH* was not detected in LCR-D. Despite the smaller size of LCR-D, it still contains a large region (250 kb) that is duplicated in LCR-A. Sequence comparison of modules from LCR-A and -D suggests that they share 97–98% nucleotide sequence identity over the entire 250 kb that is duplicated in both regions. A small region (~13 kb) in LCR-A containing marker

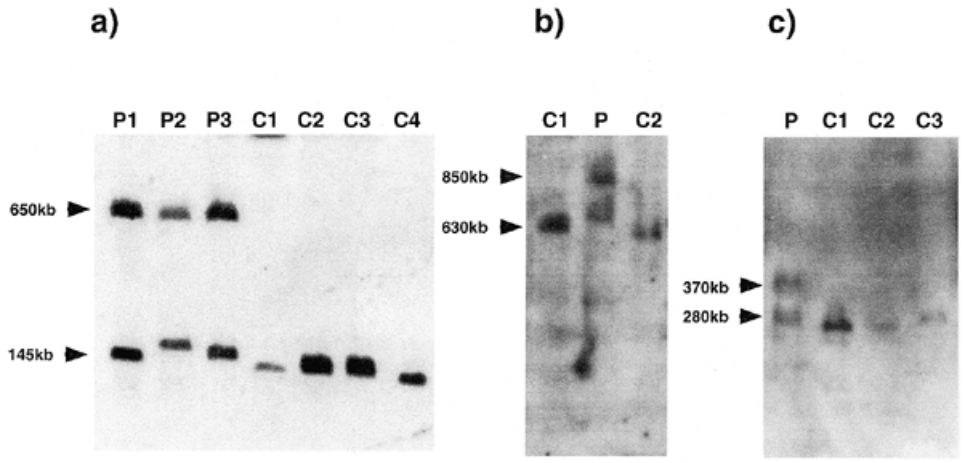

Figure 4. PFGE and Southern hybridization of genomic DNA. (**a**) *Not*I digest of A–B deletion patients (P1–P3) and non-deleted controls (C1–C4) hybridized with the *ZNF74* probe. (**b**) *Sgr*AI digest of an A–C deletion patient (P1) and non-deleted controls (C1 and C2) hybridized with the *HCF2* probe. (**c**) *Not*I digest of a C–D deletion patient (P1) and three non-deleted controls (C1–C3) hybridized with the KI 384 probe. Normal and rearranged bands are indicated by arrows with the corresponding sizes for each band.

DGCR5 (51), which is present in cosmid 46a9 and fosmid 41c7, was found to be duplicated in LCR-D in PAC 413m7 (Fig. 3A and D, pink boxes).

LCR-B is estimated to be ~135 kb. It contains one copy each of the ~45 kb module containing markers *D22S131*, VNTRL, 562f10Sp6 and *NF1L* and the ~75 kb module containing markers 444p24Sp6, *DGCR6* and *PRODH* (Fig. 3B). Interestingly, an ~15 kb module (Fig. 3B, orange box) appears to have inserted between markers VNTRL and *D22S131* within the ~45 kb module. The inserted ~15 kb module contains a duplicated copy of the *ISG43* gene (GenBank accession no. AF176642) which also is located in LCR-A (Fig. 3A, orange box). *ISG43* is an interferon-stimulated ubiquitin-specific protease (52). LCR-B shares 97–98% nucleotide sequence identity with LCR-A over the entire 135 kb that is present in both.

LCR-C is not as large as the other duplicated regions. This LCR, represented by BAC 32i11 and PAC 408l11 (Fig. 3C) contains a partial copy of the ~40 kb module containing markers *BCRL*, *HMPLPL* and *E2F6L* found in LCR-A and -D. This partial module is smaller (11–12 kb) and contains only markers *BCRL* and *HMPLPL* (Fig. 3C, blue box). Additionally, PAC 408l11 contains the gene for phosphatidylinositol 4-kinase (*PI4K*) (53). This large gene (~140 kb) extends into BAC 135h6 (Fig. 2). Interestingly, ~35 kb of the 3′ end of the *PI4K* gene also appears to be present in LCR-D in BAC 445f23. The copies of the duplicated modules in LCR-C share 97–98% nucleotide sequence identity with those present in the other LCRs.

Localization of variant deletion end-points within the chromosome 22-specific LCRs

To further examine the role of the chromosome 22-specific LCRs in deletion formation, genomic DNA from A–B, A–C and C–D deletion patients was analyzed by digestion with rare-cutting restriction enzymes and subsequent pulsed-field gel electrophoresis (PFGE). These deletions were studied because each type involves one of the smaller internal LCRs for which detailed restriction map data were readily available. After PFGE, the DNA was transferred to a nylon membrane and hybridized with unique probes immediately flanking the respective LCRs expected to contain the deletion end-points. Genomic DNA from three patients with the A–B deletion was digested with the *Not*I restriction enzyme and subjected to PFGE. Southern hybridization with a PCR-derived probe from the *ZNF74* gene, which is immediately distal to LCR-B, identifies a novel rearrangement fragment in A–B deletion patients (Fig. 4a). The *ZNF74* probe identifies a normal 145 kb fragment in all samples, including normal controls, as well as in the three A–B deletion patients. The size variation in the 145 kb normal fragment in the samples analyzed (Fig. 4a) suggests that there might be polymorphism within this region of 22q11 in the human population. In addition to the normal band, the A–B deletion patients demonstrate a novel 650 kb rearranged junction fragment created as a result of the deletion (Fig. 4a). 37g5T3, a probe immediately proximal to LCR-A, also detects the similarly sized 650 kb rearranged fragment (data not shown). These results suggest that the proximal deletion end-point of the A–B deletion is distal to 37g5T3 and that the distal end-point is proximal to *ZNF74* (Fig. 5b).

Novel junction fragments from A–C and C–D deletions have also been identified (Fig. 4b and c). Genomic DNA from an A–C deletion patient was restriction digested with enzyme *Sgr*AI followed by PFGE. Southern hybridization with a unique probe for *HCF2*, immediately distal to LCR-C (Fig. 2), identifies a normal 630 kb fragment in all samples. This fragment is seen in normal controls as well as in the A–C deletion patient (Fig. 4b). In addition to the normal band there is a novel 850 kb junction fragment in the A–C deletion patient. This suggests that the A–C deletion distal end-point lies within LCR-C. In a similar experiment genomic DNA from the C–D deletion patient (27) was digested with *Not*I and hybridized to a unique probe from KI384, immediately proximal to LCR-C (Fig. 2). In addition to a normal fragment of 280 kb, a novel 370 kb junction fragment was detected in the C–D deletion patient (Fig. 4c). This suggests that the proximal end-point of the C–D deletion resides within LCR-C.

To further sublocalize the deletion end-points of the A–B deletions, the regions of the gel containing the 145 kb normal and the 650 kb rearranged fragment were isolated for additional PCR analysis. Both gel fragments were positive for *ZNF74* by PCR (Fig. 5a), confirming the Southern hybridization data. A copy of

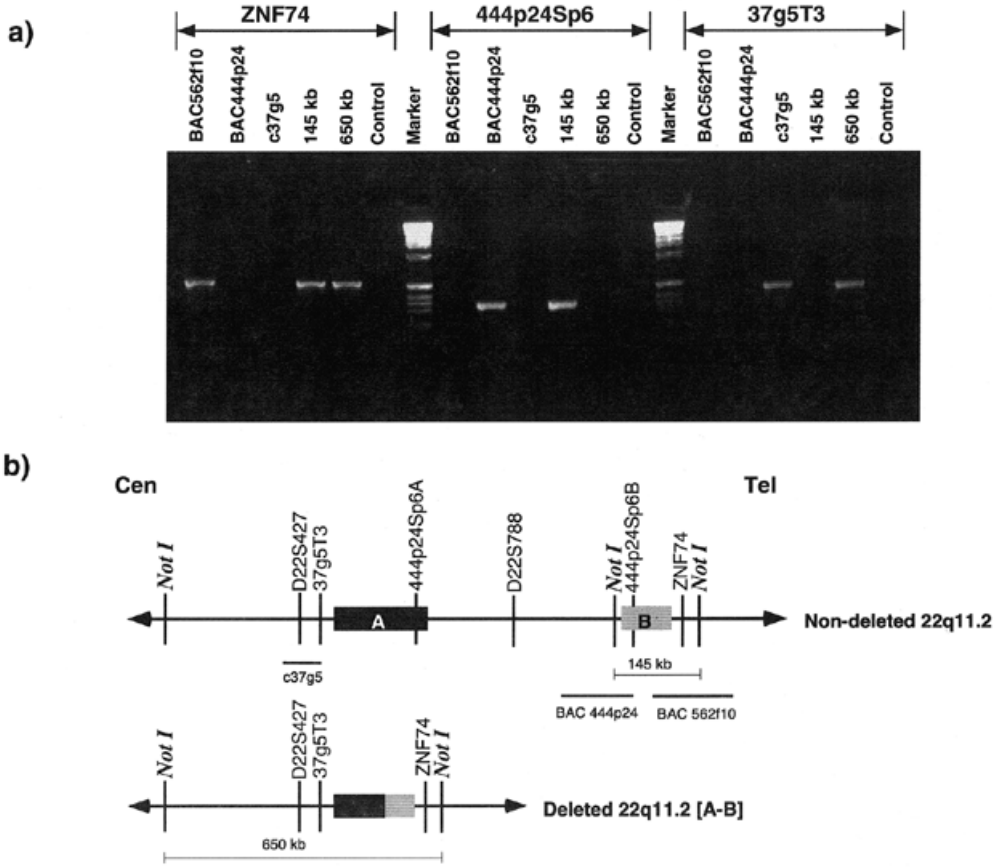

Figure 5. PCR analysis of rearranged junction fragment. The normal and rearrangement fragments from an A–B deletion patient were analyzed by PCR with different markers. (**a**) Gel electrophoresis results of PCR are shown. Each lane is labeled to indicate the template DNA tested by PCR. 145 kb, the 145 kb normal fragment; 650 kb, the 650 kb rearrangement fragment; Control, no template DNA; Marker, 1 kb DNA size marker. Each PCR marker that was tested is indicated above the lanes. All six samples were tested for *ZNF74*, 444p24Sp6 and 37g5T3. (**b**) A line illustration of a non-deleted 22q11.2 and an A–B deleted 22q11.2. The *Not*I sites that result in the 145 kb normal fragment in the non-deleted 22q11.2 and the 650 kb rearrangement fragment in the A–B deleted 22q11.2 are shown. LCR-A and -B are indicated as a black and gray box, respectively. The fusion of LCR-A and -B after the deletion is indicated by a partially black and gray box. Relevant chromosome 22 markers are shown as horizontal bars and labeled. Locations of BACs 444p24 and 562f10 are also indicated.

the PCR marker 444p24Sp6 is present within BAC 48m11 in LCR-A (Fig. 3A) as well as BAC 444p24 in LCR-B (Fig. 3B). These two copies of 444p24Sp6 have been designated as 444p24Sp6-A and -B, respectively (Fig. 5b). They can be differentiated from each other by sequence variation as they share only 98% sequence identity. 444p24Sp6 can be amplified from the 145 kb band but not from the rearranged band (Fig. 5a). Sequence of the PCR-amplified 444p24Sp6 from the 145 kb normal *ZNF74*-containing band confirmed that it corresponds to 444p24Sp6-B. Absence of this PCR product from the rearranged band suggests that both copies are absent from the junction fragment. This suggests that the distal A–B deletion end-point is within LCR-B between markers 444p24Sp6-B and *ZNF74* (Fig. 5b). Further, the absence of 444p24Sp6 and the presence of 37g5T3 in the rearranged band (Fig. 5a) localizes the proximal A–B deletion end-point within LCR-A between 37g5T3 and 444p24Sp6-A (Fig. 5b). These data localize the A–B deletion end-points within LCR-A and -B and suggest their partial deletion. Similarly, based on the sequence and Southern hybridization data, the deletion end-points of the A–C and C–D deletion appear to localize within the LCRs flanking the respective deletions (data not shown).

Evidence of duplications in non-human primates

Comparative mapping of the mouse genome in the region of conserved synteny with human 22q11.2 has revealed no evidence for duplicated blocks of sequence similar to the chromosome 22-specific LCRs (54–57). To determine whether the duplications on human chromosome 22 are human-specific, FISH analysis was performed using BAC 48m11 on metaphase spreads and interphase nuclei from three non-human primates. BAC 48m11 is from within LCR-A and contains duplicated markers *GGTL*, *E2F6L*, *HMPLPL*, *BCRL*, *D22S131*, VNTRL, 562f10sp6, *NF1L*, 444p24Sp6, *DGCR6* and *PRODH* (Figs 2 and 3). We tested two great apes, the pygmy chimpanzee (*Pan paniscus*) and the gorilla (*Gorilla gorilla*), and one Old World monkey, the rhesus monkey (*Macaca mulatta*). In humans, metaphase FISH with BAC48m11 results in one distinct signal on chromosome 22 (Fig. 6a). This is most likely because the LCRs on chromosome 22 are too close to one another to be resolved on metaphase chromosomes. Similar to what is seen in humans, BAC48m11 appeared as a single signal on the chromosomal regions orthologous to human chromosome 22 in metaphase spreads from the three non-human primates tested (Fig. 6b–d). The chromosome to which the BAC hybridized was determined by the banding pattern of DAPI counterstained images inverted to grayscale (data not shown). Therefore, in both

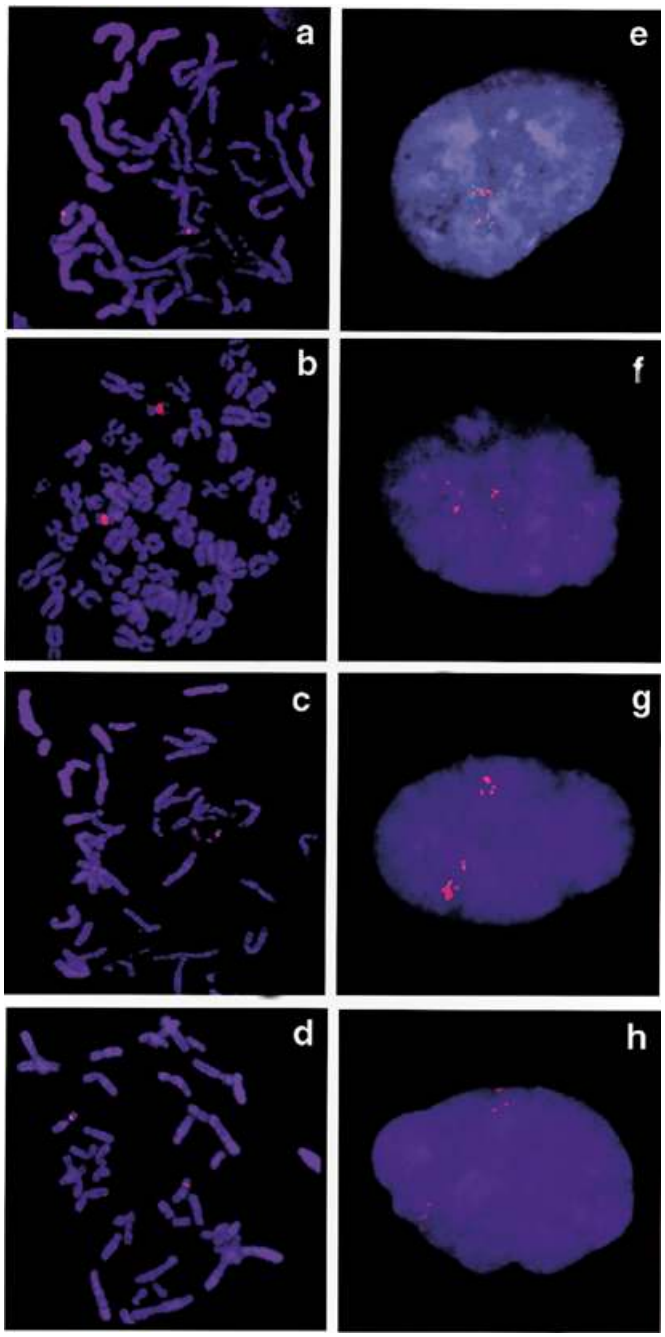

Figure 6. Evolutionary conservation of the duplication in primates. Metaphase spreads (left) and interphase nuclei (right) from various primates hybridized with BAC 48m11 labeled with digoxygenin and detected with rhodamine (red signal). (**a** and **e**) Human; (**b** and **f**) pygmy chimpanzee (*P.paniscus*); (**c** and **g**) gorilla (*G.gorilla*); (**d** and **h**) rhesus monkey (*M.mulatta*). The green signal (fluorescein) in (e) is from probe c16e8, a unique chromosome 22 cosmid probe containing marker *D22S66* (ph160b).

the chimpanzee (Fig. 6b) and gorilla (Fig. 6c), BAC48m11 signal was detected on chromosome 23, orthologous to human chromosome 22 (58,59). In the rhesus monkey, BAC48m11 signal was detected on chromosome 13 (Fig. 6d), orthologous to human chromosome 22 (60).

Since BAC48m11 appeared as a single, distinct signal when hybridized to metaphase chromosomes, it could not be determined whether this BAC contained duplicated sequence elements. Therefore, BAC 48m11 was used as a probe for FISH on G_1-arrested interphase nuclei from the different primate species. Multiple signals were observed on interphase nuclei from humans (Fig. 6e) demonstrating that the genomic DNA contained within BAC48m11 is in fact duplicated in multiple locations on chromosome 22. Similarly, multiple signals also were observed using BAC48m11 as probe on interphase nuclei from the chimpanzee, gorilla and rhesus monkey (Fig. 6f–h). This suggests that the human chromosome 22 duplication events predate the divergence of the great apes from the Old World monkeys, which is estimated to have been at least 20–25 million years ago (61).

DISCUSSION

Chromosome-specific sequence duplications have now been implicated in a number of disorders that are associated with recurrent chromosomal rearrangements. Non-random chromosomal rearrangements such as deletions, duplications and inversions have been proposed to be the result of recombination between copies of highly homologous chromosome-specific duplicons or low copy repeats. Some of the best studied examples of chromosome-specific duplicon-mediated rearrangements include Charcot–Marie–Tooth disease type 1A (CMT1A) on 17p11.2 (62,63), Prader–Willi/Angelman syndromes on 15q11–q13 (64,65), Williams–Beuren syndrome on 7q11.23 (66) and Smith–Magenis syndrome on 17p11.2 (67). The identification of chromosome 22-specific duplications, which have been referred to as LCRs, in close proximity to the 22q11.2 microdeletion end-points has led to the proposal that the LCRs somehow act to mediate the deletion event (23,24,33–36).

We have sequenced an ~3 Mb BAC/PAC/cosmid contig that represents the region commonly deleted in the 22q11.2 microdeletion associated with DGS and VCFS. Analysis of the resulting sequence data confirms the previous reports of LCRs close to the 22q11.2 deletion end-points (23,24,36,39). The LCRs share 97–98% nucleotide sequence identity within the duplicated modules which they have in common with one another. The sizes of the LCRs or the overall similarity of modular components shared among them appear to correlate with the frequency of each deletion. Therefore, the 3 Mb deletion (A–D) that is mediated by the largest of the LCRs is the most frequent (87%). These two LCRs share 250 kb of duplicated sequence in a complex arrangement of similar modules.

The 1.5 Mb (A–B) deletion which involves a common 135 kb duplicated block is the most frequent of the variant deletions, accounting for 8% of the total 22q11.2 deletions in our cohort. Interestingly, the identification of LCR-C appears to provide an explanation for the 1.5–2 Mb variant deletion being a combination of two distinct deletion populations. Therefore, a 1.5–2 Mb deletion could be either an A–B or an A–C deletion. The A–C and DGCR5-mediated variant deletions, although recurrent, are not as prevalent. Presumably this is because both of these deletions are mediated by very small duplicated blocks (<15 kb) within the LCRs. The unique deletions, including the C–D deletion which has been reported only once (27), also appear to be mediated by smaller duplicated blocks and are quite rare. Two unique, distal, variant deletions have been reported (31,32). It is likely that they also belong to this class of rare deletions mediated by smaller duplicated blocks.

Different models have been proposed to explain duplicated sequence-mediated chromosomal rearrangements associated with chromosome 22 (24). Completed sequence data for the LCRs described in this report should permit the testing of these models. Interestingly, our analysis of sequence data demonstrates that the duplicated modules that are proposed to mediate the different deletions are either in direct or inverted orientation with respect to each other within and between LCRs (Fig. 3). Therefore, there are two possible models for the formation of the deletions. The first model would involve an interchromosomal misalignment during meiosis I between the two homologs of chromosome 22. This misalignment might be mediated by the modules within separated LCRs that are in direct orientation with respect to each other. Subsequent crossing-over would lead to reciprocal deletion and duplication events. A similar mechanism causes the reciprocal duplication and deletion events that lead to Charcot–Marie–Tooth syndrome type 1A and hereditary neuropathy with liability to pressure palsies, respectively (63).

The second model to explain the formation of the deletions would involve intrachromosomal recombination between the duplicated modules. In this model the duplicated modules in inverse orientation with respect to one another might form a 'stem–loop' intermediate. Recombination between the duplicated modules forming the 'stem' would then lead to the deletion of intervening DNA present within the 'loop'. This type of rearrangement has been observed in V(D)J recombination required for the formation of immunoglobulin heavy and light chains (68). Both inter- and intra-chromosomal recombination events for the standard 3 Mb deletion have been reported (24,69). It will be of interest to determine the frequency of each type. Further, it will be important to determine the mechanism responsible for the unique deletions (26,28–32), in order to determine whether other structural changes accompany these deletions.

As described in this report, the gel isolation and PCR analysis of rearrangement fragments from multiple variant 22q11.2 deletions have further localized the breakpoints within the LCRs. The sequence of these LCRs should permit identification of the exact sites at which the chromosomal rearrangements occur in these variant 22q11.2 deletion patients. Although it is too early to determine precisely, it is of interest that the three A–B variant deletions examined appear to have similarly sized junction fragments. The isolation and analysis of the junction fragments should permit a closer examination of the issue of breakpoint clustering. It then should be possible to determine whether the breakpoints of the A–D deletion localize to specific sites or whether they are widely dispersed within LCR-A and -D. It is possible that there may be recombination hotspots within the LCRs similar to those seen in CMT1A-associated rearrangements (70). In support of this possibility, we have identified multiple copies of sequences shown to be involved in rearrangements of chromosome 22 as well as of other chromosomes within the various LCRs on 22q11.2. They include the AT-rich sequence involved in the t(17;22) in a family with neurofibromatosis type 1 (46) and the VNTR-like repeats described here. Recurrent chromosomal rearrangements shown to be mediated by VNTR-like repeats include the 1.9 Mb deletion on Xp22.3 leading to X-linked ichthyosis due to steroid sulfatase deficiency (71,72). It is possible that the presence of these highly unstable sequences may provide hotspots for chromosomal breakage within the LCRs on 22q11.2.

Once the mechanism(s) of deletion and the sequences within the LCRs that are directly involved in the 22q11.2 deletion have been identified, it will be possible to test their involvement in other chromosome 22-associated rearrangements. Chromosome 22-specific duplicated sequences already have been implicated in other recurrent constitutional rearrangements associated with this chromosome. The chromosome 22 breakpoint of the recurrent, constitutional t(11;22) has been localized to LCR-B which also is the distal end-point of the 1.5 Mb A–B variant deletion (39,43). In CES a bisatellited chromosome resulting from an inverted duplication of proximal 22q11 is present as a supernumerary chromosome (38,73). There are two distinct duplications in CES patients and their breakpoints appear to localize to LCR-A and -D, which are the proximal and distal deletion end-points, respectively, of the 3 Mb common deletion (38).

Detailed analysis of the LCRs has identified sequences homologous to genes from chromosome 22 as well as other chromosomes. The genes/pseudogenes from chromosome 22 include: *BCR-like* (*BCRL*) sequences homologous to the 3′ end of the *BCR* gene, which is involved in the t(9;22) which leads to the Philadelphia chromosome (74); *GGT-like* (*GGTL*) sequences which are homologous to the *GGT* (γ-glutamyl transpeptidase) gene (75); and *GGT-rel-like* (*GGTrelL*) which is a truncated piece of the *GGT-rel* gene (45). The genes/pseudogenes that map to other chromosomes include *HMPLP-like* (*HMPLPL*) sequences which share homology with sequences from human chromosome 5. The *HMPLP* (human membrane protein-like protein) gene shares homology with a gene for a membrane protein in rats (76). There are also *E2F6-like* (*E2F6L*) sequences homologous to a member of a family of transcription factors (47) and *V7rel* sequences (36), both of which also map to chromosome 1. Although many of the gene-related sequences present within the LCRs appear to be either non-functional pseudogenes or truncated pieces of the original genes there is evidence for several functional genes within the LCRs. These include the genes *DGCR6* (56), proline dehydrogenase (*PRODH*) (50) and *ISG43* (52). It remains to be determined whether any of the functional genes within the LCRs are copy number-sensitive which, if deleted, could be implicated in the etiology of the DGS/VCFS phenotype.

Interchromosomal duplications of a gene-rich cluster, including the *ALD* gene from Xq28 to 2p11, 10p11, 16p11 and 22q11, have been reported previously (77,78). It has been determined that the only active *ALD* gene is located on Xq28 and the duplicated copies of *ALD* on 2p11, 10p11, 16p11 and 22q11 are non-functional (78). Therefore, the *ALD* duplicons are considered to be truncated non-processed pseudogenes with little functional significance. This may be the case for the vast majority of the pseudogenes that have been identified within the chromosome 22-specific LCRs. Interestingly, the *ALD* duplications are targeted to the pericentromeric regions of other chromosomes and have been attributed to a pericentromeric-biased transposition mechanism (77–79). It is conceivable that the chromosome 22-specific LCRs may have originated by a similar mechanism involving pericentromeric duplications that expanded in size and were amplified further during primate evolution. In support of this hypothesis, FISH analysis using a duplicated sequence-containing BAC as a probe resulted in a signal near the centromere in non-human primates.

Duplications were identified in non-human primates including the chimpanzee (*P.paniscus*), the gorilla (*G.gorilla*) and the rhesus monkey (*M.mulatta*) on orthologs of human chromosome 22.

These duplications contain elements from within the chromosome 22-specific LCRs. Comparative mapping and sequence analysis of the mouse genome in the region of conserved synteny with human 22q11 has not indicated the presence of similar duplicated sequences in the rodent genome (54–57). This suggests that the chromosome 22 duplicated sequences may have evolved and amplified after the divergence of primates from rodents but before the divergence of the great apes from the Old World monkeys. Additional molecular analysis of the duplicated regions in non-human primates and comparison with the human sequences should allow us to trace the evolutionary origin and mechanisms involved in their mobility and expansion in the primate genome. It will be interesting to determine whether there is a variation in the number and organization of the duplicated regions between humans and non-human primates. If such variation exists it will provide an additional rationale to test for variation within the human population. It is tempting to speculate that there may be variation or polymorphism within the human population with respect to the number and organization of these chromosome 22-specific LCRs. Thus, there may be a critical copy number or organizational configuration of the duplicated sequences that may predispose individuals carrying them to the constitutional rearrangements associated with human chromosome 22.

MATERIALS AND METHODS

FISH analysis

For the human samples, metaphase spreads were prepared either from peripheral blood lymphocytes or lymphoblastoid cell lines using standard methodology. The pygmy chimpanzee (*P.paniscus*) (AG05253), gorilla (*G.gorilla*) (AG05251A) and rhesus monkey (*M.mulatta*) (AG08316) metaphase spreads were prepared from fibroblast cell lines obtained from Coriell Mutant Cell Repositories (Camden, NJ). Interphase nuclei were prepared according to the method described by Trask *et al.* (80). FISH was performed as previously described (81). Chromosomes were visualized by counterstaining with DAPI. Probes used for FISH were labeled by nick translation with either biotin-16 dUTP or digoxygenin-11 dUTP as described by Lichter *et al.* (82) with minor modifications. The labeled, hybridized probes were then detected by either fluorescein-conjugated avidin or rhodamine-conjugated anti-digoxygenin, respectively. The chromosome 22 FISH probes were cosmids isolated from the LL22NC03 cosmid library. BAC 48m11 which was used for FISH was from the CITB-978SK (Caltech A) BAC library (Research Genetics, Huntsville, AL).

Contig construction

Clones used to construct the contig were obtained by screening high density gridded membranes containing either BAC, PAC or cosmid libraries with radiolabeled PCR-derived probes from known markers on chromosome 22. The BAC libraries screened include the RPCI11 segment 2 from Roswell Park Cancer Institute (Buffalo, NY) and the CITB-978SK Caltech A BAC library. The PAC library screened was the RPCI3 from Roswell Park Cancer Institute and the cosmid library used was the LL22NC03 cosmid library. The filters were prehybridized and hybridized as described previously (43).

DNA sequencing and sequence analysis

End-sequencing of clones and sequencing of PCR products was performed by the Core Facility at the Children's Hospital of Philadelphia on an ABI 377 automated DNA sequencer. Cloned DNAs were isolated using Qiagen (Valencia, CA) kits and were sequenced with primers from within the vector sequence. PCR products were purified with Qiaquick (Qiagen) kits and directly sequenced with the same primers used in the PCR. The entire contig between markers *D22S427* and *D22S801* has been sequenced at the University of Oklahoma Advanced Center for Genome Technology (http://www.genome.ou.edu). Sequences of clones within the contig were obtained from the htgs data library of GenBank (http://www.ncbi.nlm.nih.gov). All sequences were masked for repeated DNA elements using the RepeatMasker web server (http://ftp.genome.washington.edu/cgi-bin/RepeatMasker ; A.F.A. Smit and P. Green, unpublished data). Masked sequences were analyzed further by BLAST searches against the GenBank database (83) to identify regions that were duplicated. Sequences from different copies of the duplicated blocks were aligned and compared to each other using ClustalW (84).

PFGE analysis and Southern hybridization

Agarose plugs from lymphoblastoid cell lines of patients were prepared in 1% low melting point agarose using standard methodology. Plugs containing DNA were digested with either *Not*I or *Sgr*AI. The digested plugs were subjected to PFGE on 1% agarose/0.5× TBE gels using a Chef Mapper (Bio-Rad, Hercules, CA). After the completion of PFGE the gels were transferred to nylon membranes and Southern hybridization was performed using standard methodology.

Extraction of DNA from PFGE gel

Preparative PFGE for DNA extraction was performed in low melting point agarose gels. The areas corresponding to both normal and rearranged bands were excised and DNAs were extracted using Gelase (Epicentre Technology, Madison, WI).

Probe generation and PCR

ZNF74, *HCF2* and KI384 probes were amplified by PCR of clones containing the respective markers. The primers used for probe synthesis were:
ZNF-F, 5′-TGCGCGAAATAGGCGCAAACT-3′, and
ZNF-R, 5′-TGAACTGTGGACAGGA CCCTC-3′, for *ZNF74*;
HCF2-F, 5′-TAGCACCATTCTTGATGTCC-3′, and
HCF2-R, 5′-CTCTAGTATGGGAGACATGG-3′, for *HCF2*;
KI384-F, 5′-AGTTTGGAATTTGCACGTCC-3′, and
KI-384-R, 5′-CCAGTTCCACCCTCTGTTGT-3′, for KI384.
For Southern hybridization, the aforementioned PCR products were radiolabeled with ^{32}P by the random priming labeling technique. The primers used to amplify the 444p24SP6 end were:
444p24SP6-F, 5′-GAAGCTGGTTCAGGTCAGAC-3′, and
444p24SP6-R, 5′-CTGGGCTTGGTCACTGTCAC-3′;
while those for the 37g5T3 end were:
37g5T3-F, 5′-CCGATCTGGAATTGAAACTC-3′, and
37g5T3-R, 5′-GCCTTTGTGCA TTGGTATGT-3′.
PCR conditions for all markers were as follows. PCR was performed in a 25 µl reaction volume containing 1× PCR buffer containing 1.5 mM $MgCl_2$ (Boehringer Mannheim, Indianapolis, IN), 1 µM of each primer, 200 µM of dNTPs and 3 U of *Taq*

polymerase (Boehringer Mannheim). PCR was performed with 5 min denaturation at 95°C followed by 35 cycles of denaturation at 95°C for 30 s, annealing at 55°C for 30 s and extension at 72°C for 1 min.

ACKNOWLEDGEMENTS

We would like to thank Bea Sellinger, Holly Mensch, Heather Mitchell, Feng Chen, Hua Qin Pan, Axin Hua and Zhili Wang for excellent technical assistance. The authors would also like to acknowledge Helma Van Grevenstein for her contribution to the work presented. The chromosome-specific gene library LL22NC03 used here was constructed at the Biomedical Sciences Division, Lawrence Livermore National Laboratory under the auspices of the National Laboratory Gene Library Project sponsored by the US Department of Energy. This work was supported by grants CA39926 and DC02027 (M.L.B. and B.S.E.), HD26979 (B.S.E.), NIH/NHGRI grant HG00313 (B.A.R.) and funds from the Charles E.H. Upham endowed chair from the Children's Hospital of Philadelphia and the University of Philadelphia (B.S.E.).

REFERENCES

1. John, B. and Freeman, M. (1975) Causes and consequences of Robertsonian exchange. *Chromosoma* **52**, 123–136.
2. Fraccaro, M., Lindsten, J., Ford, C.E. and Iselius, L. (1980) The 11q;22q translocation: a European collaborative analysis of 43 cases. *Hum. Genet.*, **56**, 21–51.
3. Zackai, E.H. and Emanuel, B.S. (1980) Site-specific reciprocal translocation, t(11;22)(q23;q11) in several unrelated families with 3:1 meiotic disjunction. *Am. J. Med. Genet.*, **7**, 507–521.
4. Kaiser, P. (1984) Pericentric inversions. Problems and significance for clinical genetics. *Hum. Genet.*, **68**, 1–47.
5. Ledbetter, D.H., Mascarello, J.T., Riccardi, V.M., Harper, V.D., Airhart, S.D. and Strobel, R.J. (1982) Chromosome 15 abnormalities and the Prader–Willi syndrome. A follow up report of 40 cases. *Am. J. Hum. Genet.*, **34**, 278–285.
6. Summar, M.L., Dasouki, M.J., Butler, M.J. and Dev, D.G. (1995) Recurring *de novo* unbalanced translocations involving 22q11.2, a hot spot for rearrangement? *Am. J. Hum. Genet.*, **57**, A128.
7. Li, M., Budarf, M.L., Chien, P., Barnoski, B.L., Emanuel, B.S. and Driscoll, D.A. (1995) Clustering of DiGeorge/velocardiofacial-associated translocations suggestive of a translocation 'hot spot'. *Am. J. Hum. Genet.*, **57**, A119.
8. Morton, N.E. (1991) Parameters of the human genome. *Proc. Natl Acad. Sci. USA*, **88**, 7474–7476.
9. Kaplan, J.C., Aurias, A., Julier, C. and Prieur, M. (1987) Human chromosome 22. *J. Med. Genet.*, **24**, 65–78.
10. McDermid, H.E., Duncan, A.M.V., Brasch, K.R., Holden, J.J.A., Magenis, E. *et al*. (1986) Characterization of the supernumerary chromosome in cat eye syndrome. *Science*, **232**, 646–648.
11. Iselius, L., Lindsten, J., Aurias, A. *et al*. (1983) The 11q;22q translocation: a collaborative study of 20 new cases and analysis of 110 families. *Hum. Genet.*, **64**, 343–355.
12. De La Chapelle, A., Herva, R., Koivisto, M. and Aula, P. (1981) A deletion in chromosome 22 can cause DiGeorge syndrome. *Hum. Genet.*, **57**, 253–256.
13. Kelley, R.I., Zackai, E.H., Emanuel, B.S., Kistenmacher, M., Greenberg, F. and Punnett, H.H. (1982) The association of the DiGeorge anomalad with partial monosomy of chromosome 22. *J. Pediatr.*, **101**, 197–200.
14. Driscoll, D.A., Budarf, M.L. and Emanuel, B.S. (1992) A genetic etiology for DiGeorge syndrome: consistent deletions and microdeletions of 22q11. *Am. J. Hum. Genet.*, **50**, 924–933.
15. Driscoll, D.A., Spinner, N.B., Budarf, M.L., McDonald-McGinn, D.M., Zackai, E.H. *et al.* (1992) Deletions and microdeletions of 22q11.2 in velo-cardio-facial syndrome. *Am. J. Med. Genet.*, **44**, 261–268.
16. Carey, A.H., Kelly, D., Halford, S., Wadey, R., Wilson, D., Goodship, J., Burn, J., Paul, T., Sharkey, A., Dumanski, J. *et al.* (1992) Molecular genetic study of the frequency of monosomy 22q11 in DiGeorge syndrome. *Am. J. Hum. Genet.*, **51**, 964–970.
17. Driscoll, D.A., Salvin, J., Sellinger, B., Budarf, M.L., McDonald-McGinn, D.M., Zackai, E.H. and Emanuel, B.S. (1993) Prevalence of 22q11 microdeletions in DiGeorge and velocardiofacial syndromes: implications for genetic counselling and prenatal diagnosis. *J. Med. Genet.*, **30**, 813–817.
18. Burn, J., Takao, A., Wilson, D., Cross, I., Momma, K., Wadey, R., Scambler, P. and Goodship, J. (1993) Conotruncal anomaly face syndrome is associated with a deletion within chromosome 22. *J. Med. Genet.*, **30**, 822–824.
19. Kelley, D., Goldberg, R., Wilson, D., Lindsay, E., Carey, A., Goodship, J., Burn, J., Cross, I., Shprintzen, R.J. and Scambler, P.J. (1993) Confirmation that the velo-cardiofacial syndrome is associated with haplo-insufficiency of genes at chromosome 22. *Am. J. Med. Genet.*, **45**, 308–312.
20. Burn, J. and Goodship, J. (1996) Congenital heart disease. In Rimoin, D.L., Conner, J.M., Pyeritz, R.E. and Emery, A.E.H. (eds), *Emery and Rimoin's Principles and Practice of Medical Genetics*, Vol. 1. Churchill Livingstone, New York, NY, 767–803.
21. Lindsay, E.A., Greenberg, F., Shaffer, L.G., Scamber, P.J. and Baldini, A. (1995) Submicroscopic deletions at 22q11.2: variability of the clinical picture and delineation of a common deleted region. *Am. J. Med. Genet.*, **56**, 191–197.
22. Morrow, B., Goldberg, R., Carlson, C., Dasgupta, R., Sirotkin, H., Collins, J., Dunham, I. *et al.* (1995) Molecular definition of the 22q11 deletions in velo-cardio-facial syndrome. *Am. J. Hum. Genet.*, **56**, 1391–1403.
23. Emanuel, B.S., Goldmuntz, E., Budarf, M.L., Shaikh, T., McGrath, J., McDonald-McGinn, D., Zackai, E.H. *et al*. (1999) Blocks of duplicated sequence define the endpoints of DGS/VCFS 22q11.2 deletions. In Clark, E., Nakazawa, M. and Takao, A. (eds), *Etiology and Morphogenesis of Congenital Heart Disease*. Futura, Armonk, NY.
24. Edelmann, L., Pandita, R.K., Spiteri, E., Funke, B., Goldberg, R., Palaiswamy, N., Chaganti, R.S.K., Magenis, E., Shprintzen, R.J. and Morrow, B.E. (1999) A common molecular basis for rearrangement disorders on chromosome 22q11. *Hum. Mol. Genet.*, **8**, 1157–1167.
25. Carlson, C., Sirotkin, H., Pandita, R., Goldberg, R., Mckie, J., Wadey, R., Patanjali, S.R. *et al*. (1997) Molecular definition of 22q11 deletions in 151 velo-cardio-facial syndrome patients. *Am. J. Hum. Genet.*, **61**, 620–629.
26. Levy, A., Demczuk, S., Aurias, A., Depetris, D., Mattei, M.G. and Philip, N. (1995) Interstitial 22q11 deletion excluding the ADU breakpoint in a patient with DGS. *Hum. Mol. Genet.*, **4**, 2417–2418.
27. Kurahashi, H., Tsuda, E., Kohama, R., Nakayama, T., Masuno, M., Imaizumi, K., Kamiya, T., Sano, T., Okada, S. and Nishisho, I. (1997) Another critical region for deletion of 22q11: a study of 100 patients. *Am. J. Med. Genet.*, **72**, 180–185.
28. O'Donnell, H., McKeown, C., Gould, C., Morrow, B. and Scambler, P. (1997) Detection of a deletion within 22q11 which has no overlap with the DiGeorge syndrome critical region. *Am. J. Hum. Genet.*, **60**, 1544–1548.
29. Yamagishi, H., Garg, V., Matsuoka, R., Thomas, T. and Srivastava, D. (1999) A molecular pathway revealing a genetic basis for human cardiac and craniofacial defects. *Science*, **283**, 1158–1161.
30. McQuade, L., Christodoulou, J., Budarf, M.L., Sachdev, R., Wilson, M., Emanuel, B. and Colley, A. (1999) A further case of a 22q11.2 deletion with no overlap of the minimal DiGeorge Syndrome critical region (MDGCR). The involvement of the TBX1 and COMT genes in the deletion and the clinical phenotype. *Am. J. Med. Genet.*, **86**, 27–33.
31. Rauch, A., Pfeiffer, R.A., Leipold, G., Singer, H., Tigges, M. and Hofbeck, M. (1999) A novel 22q11.2 microdeletion in DiGeorge Syndrome. *Am. J. Hum. Genet.*, **64**, 659–667.
32. Saitta, S.C., McGrath, J.M., Mensch, H., Shaikh, T.H., Zackai, E.H. and Emanuel, B.S. (1999) A 22q11.2 deletion which excludes UFD1L and CDC45L in a patient with conotruncal and craniofacial defects. *Am. J. Hum. Genet.*, **65**, 562–566.
33. Halford, S., Lindsay, E., Natudu, M., Carey, A.H., Baldini, A. and Scambler, P.J. (1993) Low-copy-number repeat sequences flank the DiGeorge/velo-cardio-facial syndrome loci at 22q11. *Hum. Mol. Genet.*, **2**, 191–196.
34. Lindsay, E.A., Halford, S., Wadey, R., Scambler, P.J. and Baldini, A. (1993) Molecular cytogenetic characterization of the DiGeorge syndrome region using fluorescence *in situ* hybridization. *Genomics*, **17**, 403–407.
35. Collins, J.E., Mungall, A.J., Badcock, K.L., Fay, J.M. and Dunham, I. (1997) The organization of gamma-glutamyltransferase genes and other low copy repeats in human chromosome 22q11. *Genome Res.*, **7**, 522–531.

36. Edelmann, L., Pandita, R.K. and Morrow, B.E. (1999) Low-copy repeats mediate the common 3-Mb deletion in patients with velo-cardio-facial syndrome. *Am. J. Hum. Genet.*, **64**, 1076–1086.
37. Emanuel, B.S., Budarf, M.L., Shaikh, T. and Driscoll, D.A. (1998) Blocks of duplicated sequence define the endpoints of DGS/VCFS 22q11.2 deletions. *Am. J. Hum. Genet.*, **63**, A11.
38. McTaggart, K.E., Budarf, M.L., Driscoll, D.A., Emanuel, B.S., Ferreira, P. and McDermid, H.E. (1998) Cat eye syndrome chromosome breakpoint clustering: identification of two intervals also associated with 22q11 deletion syndrome breakpoints. *Cytogenet. Cell Genet.*, **81**, 222–228.
39. Funke, B., Edelmann, L., McCain, N., Pandita, R.K., Ferreira, J., Merscher, S., Zohouri, M. *et al.* (1999) Der(22) syndrome and Velo-Cardio-Facial syndrome/DiGeorge syndrome share a 1.5-Mb region of overlap on chromosome 22q11. *Am. J. Hum. Genet.*, **64**, 747–758.
40. Dunham, I., Shimizu, N., Roe, B.A., Chissoe, S., Hunt, A.R., Collins, J.E., Bruskiewich, R., Beare, D.M., Clamp, M., Smink, L.J. *et al.* (1999) The DNA sequence of human chromosome 22. *Nature*, **402**, 489–495.
41. Emanuel, B.S., Budarf, M.L. and Scambler, P.J. (1999) The genetic basis of conotruncal cardiac defects: the chromosome 22q11.2 deletion. In Harvey, R. and Rosenthal, N. (eds), *Heart Development*. Academic Press, New York, NY, 463–478.
42. Budarf, M.L., Eckman, B., Michaud, D., McDonald, T., Gavigan, S., Buetow, K.H., Tatsumura, Y., Lui, Z., Hilliard, C., Driscoll, D. *et al.* (1996) Regional localization of over 300 loci on human chromosome 22 using a somatic cell hybrid mapping panel. *Genomics*, **35**, 275–288.
43. Shaikh, T.H., Budarf, M.L., Celle, L., Zackai, E.H. and Emanuel, B.S. (1999) Clustered 11q23 and 22q11 breakpoints and 3:1 meiotic malsegregation in multiple unrelated t(11;22) families. *Am. J. Hum. Genet.*, **65**, 1595–1607.
44. Budarf, M., Canaani, E. and Emanuel, B.S. (1988) Linear order of the four BCR-related loci in 22q11. *Genomics*, **3**, 168–171.
45. Heisterkamp, N., Rajpert-De Meyts, E., Uribe, L., Forman, H.J. and Groffen, J. (1991) Identification of a human gamma-glutamyl cleaving enzyme related to, but distinct from, gamma-glutamyl transpeptidase. *Proc. Natl Acad. Sci. USA*, **88**, 6303–6307.
46. Kehrer-Sawatzki, H., Haussler, J., Krone, W., Bode, H., Jenne, D.E., Mehnert, K.U., Tummers, U. and Assum, G. (1997) The second case of a t(17;22) in a family with neurofibromatosis type 1: sequence analysis of the breakpoint regions. *Hum. Genet.*, **99**, 237–247.
47. Cartwright, P., Muller, H., Wagener, C., Holm, K. and Helin, K. (1998) E2F-6: a novel member of the E2F family is an inhibitor of E2F-dependent transcription. *Oncogene*, **17**, 611–623.
48. Ishikawa, K., Nagase, T., Suyama, M., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N. and Ohara, O. (1998) Prediction of the coding sequences of unidentified human genes. X. The complete sequences of 100 new cDNA clones from brain which can code for large proteins *in vitro*. *DNA Res.*, **5**, 169–176.
49. Demczuk, S., Thomas, G. and Aurias, A. (1996) Isolation of a novel gene from the DiGeorge syndrome critical region with homology to *Drosophila gdl* and to human *LAMC1* genes. *Hum. Mol. Genet.*, **5**, 633–638.
50. Gogos, J.A., Santha, M., Takacs, Z., Beck, K.D., Luine, V., Lucas, L.R., Nadler, J.V. and Karayiorgou, M. (1999) The gene encoding proline dehydrogenase modulates sensorimotor gating in mice. *Nature Genet.*, **21**, 434–439.
51. Gong, W., Emanuel, B.S., Collins, J., Kim, D.H., Wang, Z., Chen, F., Zhang, G., Roe, B. and Budarf, M.L. (1996) A transcription map of the DiGeorge and velo-cardio-facial syndrome critical region on 22q11. *Hum. Mol. Genet.*, **5**, 789–800.
52. Li, X.-L., Blackford, J.A., Judge, C.S., Liu, M., Xiao, W., Kalvakolanu, D.V. and Hassel, B.A. (1999) RNase-L-dependent regulation of a novel interferon-stimulated gene: feedback inhibition of the interferon response (GenBank accession no. AF176642).
53. Wong, K. and Cantley, L.C. (1994) Cloning and characterization of a human phosphatidylinositol 4-kinase. *J. Biol. Chem.*, **269**, 28878–28884.
54. Lund, J., Roe, B., Chen, F., Budarf, M., Galili, N., Riblet, R., Miller, R.D., Emanuel, B.S. and Reeves, R.H. (1999) Sequence-ready physical map of the mouse chromosome 16 region with conserved synteny to the human velocardiofacial syndrome region on 22q11.2. *Mamm. Genome*, **10**, 438–443.
55. Botta, A., Lindsay, E.A., Jurecic, V. and Baldini, A. (1997) Comparative mapping of the DiGeorge syndrome region in mouse shows inconsistent gene order and differential degree of gene conversion. *Mamm. Genome*, **8**, 890–895.
56. Puech, A., Saint-Jore, B., Funke, B., Gilbert, D.J., Sirotkin, H., Copeland, N.G., Jenkins, N.A., Kucherlapati, R., Morrow, B. and Skoultchi, A.I. (1997) Comparative mapping of the human 22q11 chromosomal region and the orthologous region in mice reveals complex changes in gene organization. *Proc. Natl Acad. Sci. USA*, **94**, 14608–14613.
57. Sutherland, H.F., Kim, U.-J. and Scambler, P.J. (1998) Cloning and comparative mapping of the DiGeorge syndrome critical region in the mouse. *Genomics*, **52**, 37–43.
58. Yunis, J.J. and Prakash, O. (1982) The origin of man: a chromosomal pictorial legacy. *Science*, **215**, 1525–1530.
59. Jauch, A., Weinberg, J., Stanyon, R., Arnold, N., Tofanelli, S., Ishida, T. and Cremer, T. (1992) Reconstruction of genomic rearrangements in great apes and gibbons by chromosome painiting. *Proc. Natl Acad. Sci. USA*, **89**, 8611–8615.
60. Wienberg, J., Stanyon, R., Jauch, A. and Cremer, T. (1992) Homologies in human and *Macaca fuscata* chromosomes revealed by *in situ* suppression hybridization with human chromosome specific DNA libraries. *Chromosoma*, **101**, 265–270.
61. Kumar, S. and Hedges, B.S. (1998) A molecular timescale for vertebrate evolution. *Nature*, **392**, 917–920.
62. Chance, P.F., Abbas, N., Lensch, M.W., Pentao, L., Rao, B.B., Patel, P.I., and Lupski, J.R. (1994) Two autosomal dominant neuropathies result from reciprocal DNA duplication/deletion of a region on chromosome 17. *Hum. Mol. Genet.*, **3**, 223–228.
63. Lupski, J.R. (1998) Charcot–Marie–Tooth disease: lessons in genetic mechanisms. *Mol. Med.*, **4**, 3–11.
64. Christian, S.L., Fantes, J.A., Mewborn, S.K., Huang, B. and Ledbetter, D.H. (1999) Large genomic dulicons map to sites of instability in Prader–Willi/Angelman syndrome chromosome region (15q11–q13). *Hum. Mol. Genet.*, **8**, 1025–1037.
65. Amos-Landgraf, J.M., Ji, Y., Gottlieb, W., Depinet, T., Wandstrat, A.E., Cassidy, S.B., Driscoll, D.J., Rogan, P.K., Schwartz, S. and Nicholls, R.D. (1999) Chromosome breakage in the Prader–Willi and Angelman syndromes involves recombination between large, transcribed repeats at proximal and distal breakpoints. *Am. J. Hum. Genet.*, **65**, 370–386.
66. Perez-Jurado, L.A., Wang, Y.K., Peoples, R., Coloma, A., Cruces, J. and Francke, U. (1998) A duplicated gene in the breakpoint regions of the 7q11.23 Williams–Beuren syndrome deletion encodes the initiator binding protein TFII-I and BAP-135, a phosphorylation target of BTK. *Hum. Mol. Genet.*, **7**, 325–334.
67. Chen, K.S., Manian, P., Koeuth, T., Potocki, L., Zhao, Q., Chinault, A.C., Lee, C.C. *et al.* (1997) Homologous recombination of a flanking repeat gene cluster is a mechanism for a common contiguous gene deletion syndrome. *Nature Genet.*, **17**, 154–163.
68. Lewis, S.M. and Wu, G.E. (1997) The origin of V(D)J recombination. *Cell*, **88**, 159–162.
69. Baumer, A., Dutly, F., Balmer, D., Riegel, M., Tukel, T., Krajewska-Walasek, M. and Schinzel, A.A. (1998) High level of unequal meiotic crossovers at the origin of the 22q11.2 and 7q11.23 deletions. *Hum. Mol. Genet.*, **7**, 887–894.
70. Reiter, L.T., Hastings, P.J., Nelis, E., De Jong, P., Van Broeckhoven, D. and Lupski, J. (1996) Human meiotic recombination products revealed by sequencing a hotspot for homologous strand exchange in multiple HNPP deletion patients. *Am. J. Hum. Genet.*, **62**, 1023–1033.
71. Ballabio, A., Bardoni, B., Guioli, S., Basler, E. and Camerino, G. (1990) Two families of low-copy-number repeats are interspersed on Xp22.3: implications for the high frequency of deletions in this region. *Genomics*, **8**, 263–270.
72. Yen, P.H., Li, X.M., Tsai, S.P., Johnson, C., Mohandas, T. and Shapiro, L.J. (1990) Frequent deletions of the human X chromosome distal short arm result from recombination between low copy repetitive elements. *Cell*, **61**, 603–610.
73. Mears, A.J., Duncan, A.M.V., Budarf, M.L., Emanuel, B.S., Sellinger, B., Siegel-Bartelt, J., Greenberg, C.R. and McDermid, H.E. (1994) Molecular characterization of the marker chromosome associated with Cat Eye Syndrome. *Am. J. Hum. Genet.*, **55**, 134–142.
74. Chissoe, S.L., Bodenteich, A., Wang, Y.-F., Wang Y.-P., Burian, D., Clifton, S.W., Crabtree, J., Freeman, A., Iyer, K., Jian, L. *et al.* (1995) Sequence and analysis of the human ABL gene, the BCR gene, and regions involved in the Philadelphia chromosomal translocation. *Genomics*, **27**, 67–82.
75. Morris, C., Courtay, C., van Kassel, A.G., ten Hoeve, J., Heisterkamp, N. and Groffen, J. (1993) Localization of a gamma-glutamyl-transferase-related gene family on chromosome 22. *Hum. Genet.*, **91**, 31–36.
76. Soderqvist, H., Jiang, W.Q., Ringertz, N. and Hallberg, E. (1996) Formation of nuclear bodies in cells overexpressing the nuclear pore protein POM121. *Exp. Cell Res.*, **225**, 75–84.

77. Eichler, E.E., Lu, F., Shen, Y., Antinacci, R., Jurecic, V., Doggett, N.A., Moyzis, R.K., Baldini, A., Gibbs, R.A. and Nelson, D.L. (1996) Duplication of a gene rich cluster between 16p11.1 and Xq28: a novel pericentromeric-directed mechanism for paralogous genome evolution. *Hum. Mol. Genet.*, **5**, 899–912.
78. Eichler, E.E., Budarf, M.L., Rocchi, M., Deaven, L.L., Doggett, N.A., Baldini, A., Nelson, D.L. and Mohrenweiser, H. (1997) Interchromosomal duplications of the adrenoleukodystrophy locus: a phenomenon of pericentromeric plasticity. *Hum. Mol. Genet.*, **6**, 991–1002.
79. Regnier, V., Meddeb, M., Lecointre, G., Richard, F., Duverger, A., Nguyen, V., Dutrillaux, B., Bernheim A. and Danglot, G. (1997) Emergence and scattering of multiple neurofibromatosis (NF-1)-related sequences during hominoid evolution suggest a process of pericentromeric interchromosomal transposition. *Hum. Mol. Genet.*, **6**, 9–16.
80. Trask, B.J., Massa, H., Kenwrick, S. and Gitschier, J. (1991) Mapping of human chromosome Xq28 by two-color fluorescence *in situ* hybridization of DNA sequences to interphase cell nuclei. *Am. J. Hum. Genet.*, **48**, 1–15.
81. Holmes, S.E., Riazi, M.A., Gong, W., McDermid, H.E., Sellinger, B.T., Hua, A., Chen, F. *et al.* (1997) Disruption of the clathrin heavy chain-like gene (CLTCL) associated with features of DGS/VCFS: a balanced (21;22)(p12;q11) translocation. *Hum. Mol. Genet.*, **6**, 357–367.
82. Lichter, P., Cremer, T., Borden, J., Manuelidis, L. and Ward, D.C. (1988) Delineation of individual human chromosomes in metaphase and interphase cells by *in situ* suppression hybridization using recombinant DNA libraries. *Hum. Genet.*, **80**, 224–234.
83. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.
84. Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673–4680.